Digital Video Needs for Oceanographic Images for the National Oceanic and Atmospheric Administration (NOAA): Phase 2

Dottie Anderson, NOAA Central Library
1315 East West Highway, SSMC 3 2nd Floor
Silver Spring, Maryland 20910, USA
Dorothy.Anderson@noaa.gov

Mary Lou Cumberpatch, NOAA Central Library,
1315 East West Highway, SSMC 3 2nd Floor
Silver Spring, Maryland
Mary.Lou.Cumberpatch@noaa.gov

Janice A. Beattie, NOAA Central Library
1315 East West Highway, SSMC 3 2nd Floor
Silver Spring, Maryland
Janice.Beattie@noaa.gov

Anna Fiolek, NOAA Central Library
1315 East West Highway, SSMC 3 2nd Floor
Silver Spring, Maryland
Anna.Fiolek@noaa.gov

Abstract—Existing programs within The National Oceanic and Atmospheric Administration (NOAA) for data archiving and access do not address the increasing requirements for managing digital video, imagery, and audio resources. In particular, NOAA's Office of Ocean Exploration (OE) has a need for an organized digital management system to handle increasing amounts of video, still imagery, and audio resources generated by OE programs and grants. The NOAA Central Library (NCL), in collaboration with NOAA's Office of OE, embarked on a pilot project in 2002 to explore the possibilities of providing such data for the exploration research community, as well as educators and the general public. The NCL plans to develop a standardized capability for public access to video, image, and audio information from NOAA's efforts in uncovering and documenting undersea habitats and marine species. Eventually, the goal is to set up a portal from which all OE video/image/audio will be accessible. This paper will discuss the progress of the pilot project during its first year, and present the goals for the next year. It will discuss the research, standardization process (metadata), and implementation phases. It will assess user needs and requirements, developed policies and procedures, and explored various options for a retrieval system for the video/image/audio. It will discuss many of the questions faced in the first phase of the project including: Would the library be the central repository for all of NOAA? How would the video/images/audio material be collected? How would it be accessed? How would it be archived? The ultimate goal of the project is to provide easy access to the rich and varied digital video, images and audio material generated by NOAA's Office of OE.

I. INTRODUCTION

The goals and objectives of the Office of Ocean Exploration (OE) Program in the National Oceanic and Atmospheric Administration (NOAA) are to increase the knowledge of the oceans through exploration and discovery, to share ocean discoveries through outreach to the public and educators, and to sustain NOAA's capability and leadership in ocean exploration. One of the objectives of the OE program is to catalog and document the results of ocean exploration. To do that, NOAA needs to support the establishment of a new knowledge base for a broad, multidisciplinary user community, including educators, students, researchers, commercial industry and government policy makers and the general public.

As described in the NOAA Treasures at Risk [1] report, users are demanding online ordering, searching, and browsing capabilities with electronic file transfer for data delivery. Users are no longer content to wait days or weeks to obtain their data or information. Increasingly, users want information rather than data, as information and products derived from observations are frequently more useful to the public and industry than the original data. This demand is one of the principal drivers in causing the development of new data handling and display techniques to manage ocean exploration data.

The results from the cruises will be cataloged and documented using the latest in database technology, communications, mapping techniques, and recording media. Data products will include reports, maps, inventories, geographic information system layers, and databases. The NCL has played an active role in the formation of a Video Data Management System (VDMS) Plan for the Office of OE within NOAA. The VDMS is part of the larger OE Data Management Plan. This VDMS will work for all recorded media, such as, video, stills, and audio.

To accomplish these goals and objectives, the NOAA Research Office of OE put together an Integrated Product Team Advisory Group and Integrated Product Team (IPT) for Oceanographic Data Management. The IPT Advisory Group is composed of NOAA upper management, non-OE NOAA staff involved in data management, and non-NOAA personnel who are involved in OE grants and contracts. The IPT is a working group composed of members from different offices throughout NOAA to work cooperatively on the issue of oceanographic data management particularly for the Office of OE. Ms. Janice Beattie, the director of the NOAA Central Library (NCL), is the project manager for the IPT. The mission of the IPT for oceanographic data management is to produce a data management plan. It will prepare a blueprint for metadata collection, processing, archiving, and distribution. As such, the IPT will identify specific actions, products, and other deliverables that will help meet OE needs. The IPT will develop the elements, activities, staff, budget, and schedule for each deliverable. In addition, the IPT will organize and coordinate teams to produce agreed upon deliverables.

Over the long-term, the IPT will evaluate the effectiveness of the data management plan, assist with modifying the components of the data management system, and identify deliverables to meet future needs. The IPT may also address other NOAA data management needs as appropriate.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 SEP 2003	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Digital Video Needs for Oceanographic Images for the National Oceanic and Atmospheric Administration (NOAA): Phase 2	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **NOAA Central Library 1315 East West Highway, SSMC 3 2nd Floor Silver Spring, Maryland 20910, USA**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002146. Oceans 2003 MTS/IEEE Conference, held in San Diego, California on September 22-26, 2003. U.S. Government or Federal Purpose Rights License

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**5**	

II. THE IPT WORKING GROUPS

The IPT was formed in December 2002. At the first meeting, the IPT defined its mission as developing a data management strategy for NOAA's Office of OE to ensure that data and information collected during OE-funded expeditions will be made accessible to the widest audience possible. In addition, the IPT listed potential deliverable products that would form the backbone of a comprehensive data management strategy. The deliverables were identified as a data entry system, metadata requirements, a digital atlas or encyclopedia, a video data management system, cruise reports, a still-image gallery, and a data management system for cataloging, storage, access, and archive.

The IPT created five Working Groups responsible for developing these products: the Video/Image/Audio Working Group, the Metadata Working Group, the Cruise Reports Working Group, the Data Management Process Working Group, and the Digital Atlas Working Group. The challenge, of course, was for each working group to incorporate the needs of all the working groups into their own specific tasks and deliverables.

A. The Video/Image/Audio Working Group

This group is tasked to develop a Video Data Management System for NOAA's Office of OE. The group is tasked to (1) develop requirements to archive video, still imagery, and acoustic (multimedia) data in accessible formats for search, retrieval, and access by the public, (2) coordinate with the Office of OE to ensure Principal Investigators (PIs) on expeditions have the tools to generate appropriate metadata, and (3) ensure its functional and operational requirements are addressed by the other IPT Working Groups.

B. The Metadata Working Group

This group has developed a "metadata cookbook" for use by PIs and data managers participating in OE expeditions. The purpose of this cookbook is to provide guidance on creating metadata for video, audio, acoustic, bathymetric, biological, chemical and oceanographic data generated by the expeditions. The cookbook contains required elements pertaining specifically to video, still images, and acoustic data, which will allow the NCL to provide access to the data through the NOAA Libraries Online Catalog (NOAALINC) (http://www.lib.noaa.gov) and the worldwide online library catalog, WorldCat (http://www.oclc.org/worldcat/). WorldCat is the world's largest and richest database of bibliographic information linking catalogs of worldwide libraries.

C. The Cruise Reports Working Group

This group has created a list of elements required for the metadata descriptions, as well as a template for a cover sheet. An example of a required element is "Summary of Digital Data Collected," which asks for the PI to identify volume in MB/GB/TB, etc., and identify the type of data collected. The PI should be as explicit as possible (e.g., identify high-definition video as opposed to simply video).

D. The Data Management Process Working Group

This group is tasked with developing the OE Central Catalog containing Federal Geographic Data Committee (FGDC) metadata for all OE expeditions. The OE Central Catalog is designed to allow FGDC metadata records to be crosswalked to library standard MARC 21 and exported to NOAALINC. MARC is the acronym for Machine-Readable Cataloging. It defines a data format, which emerged from a Library of Congress led initiative begun thirty years ago. MARC became USMARC in the 1980s and MARC 21 in the late 1990s. It provides the mechanism by which computers exchange, use and interpret bibliographic information and its data elements make up the foundation of most library catalogs used today. Metadata describing the cruise reports and other project reports for all OE expeditions will also be available through NOAALINC and WorldCat.

E. The Digital Atlas Working Group

This group has developed a Geographic Information System (GIS)-based Internet tool to help the public locate data and recorded media generated from each NOAA-sponsored OE expedition, including multimedia. Video, acoustic, and still images metadata will be in the Digital Atlas, the NCL Online Catalog and the OE Central Data Catalog (under development). The pilot Digital Atlas is available at http://www.ncddc.noaa.gov/OE.

III. VIDEO DATA MANAGEMENT SYSTEM

The IPT Video/Image/Audio Working Group is hard at work drafting a requirements document for the VDMS. Members of this team have come from the library staff, the National Environmental Satellite Data and Information Service (NESDIS) Visualization Laboratory, contract staff working on the project, and members from academia and other NOAA offices.

The VDMS Requirements can be divided into seven distinct activities: data collection, acquisitions, cataloging, recorded message processing, access, library archive, and user distribution.

A. Data Collection

This requirement is dependent on coordination between the NCL and the OE Data Manager. The NCL depends on the OE data manager to see that all video image files, still-image files, and audio-image files are collected. Quality control and basic recorded media editing must be performed on all media. The PI will use an inclusive onboard system to edit recorded media. Recorded media will be checked for consistency with its associated metadata, and a MARC 21 record will be produced using the manual application of the FGDC-MARC 21 Crosswalk. After these steps have been completed, the OE Catalog Manager will export the metadata record and transfer the recorded media files to the NCL.

B. Acquisitions

During this phase of the project, the NCL acquires all recorded media and performs further editing, composing, and annotation. The NCL will acquire the original generated content of all recorded media for archiving and inclusion into the library electronic catalog. A NCL liaison will ensure that both an original and a working copy of all material are located and supplied to the library. The library will locate metadata for each instance of recorded media supplied by the PI and confirm receipt of the recorded media. All recorded media and related data will be logged into a standard form with each item having a unique accession

number. Once it is confirmed that all metadata is compliant and complete, it will be sent to cataloging. Non-compliant or non-complete metadata will be sent back to the OE Catalog manager at a NOAA data center.

C. Cataloging

The NCL cataloging staff will perform the necessary electronic library cataloging functions. The NCL Metadata Librarian will integrate access points into the library catalog record to link all recorded media and OE multimedia records from each cruise. The Metadata Librarian will use the Digital Metadata 12 fields developed by library staff to ensure that all necessary information is included in the NOAALINC. The Digital Metadata 12 files are the necessary data for a complete metadata record in the library catalog. A unique catalog record will be created for each whole recorded media. The library catalog will ensure the ability to search all cruise metadata at once or for a specific item in a cruise.

D. Recorded Media Processing

Recorded Media Processing encompasses all steps taken between removing the original content from the recording device through its availability on the web, the human resources required, and hardware and software. Intermediate steps involve editing, constructing the video stream, extracting still images, annotating metadata and integrating that metadata with the web interface. An assumption is made at this time that the NCL will be the focal point for this activity.

Recorded Media Processing will receive the original content designated for inclusion in the online OE cruise database from the NCL acquisition staff. The Recorded Media Processing staff person will edit the recorded media for content and define the recorded media encoding and streaming standards.

Recorded Media Processing is committed to total open-systems architecture. There is no standardization for encoding and streaming at this time. The Moving Picture Experts Group (MPEG) standard is the farthest along in gaining universal acceptance. Recorded Media Processing will identify online media formats for users. Next, the staff person will annotate the recorded media metadata in accordance with the established OE metadata crosswalk template and MERMaid-like code interface. MERMaid (Metadata Enterprise Resource Management Aid) allows users/data providers to establish unlimited metadata databases to organize the metadata records anyway they see fit. The recorded media will then be available for posting on the Internet through dedicated streaming servers.

E. Access

Users of the NCL will be able to access the video data from several points—the NOAALINC, the OE Central Catalog, a future NOAA video portal, and WorldCat. Input will be provided from both the NOAA Data Centers and the Internet for video images. Multi-platform video image access and retrieval will be available for the NCL. There will be a gateway link from the Data Center where the OE Central Catalog will reside. The NCL will provide Z39.50 [2] mapping attributes for the OE catalog gateway and have a web link for input from other sources to merge video into other products from a graphics processor. Multi-platform access will accommodate PC or MAC users. It is important that library users be able to extract cruises at the following defined levels: access only frame grabs if desired, access streaming videos, or find the entire video image file, if necessary.

F. Library Archive

Video image files will be archived to CD-ROM, DVD, and Jukeboxes. The library will possess the physical medium including logbooks, videotapes, audiotapes, audio files, and moving picture files. The library will also possess the digital medium, including electronic documents, backup copies, and archival copies. There will be a full system backup of each video image file and a copy stored off site at another NOAA library or to-be-determined site.

G. User Distribution

The NCL, in collaboration with the OE Community, will promote NOAA's OE program to academia, researchers, students, and teachers of levels K–12. Staff members will promote NOAA's Ocean Exploration program at conferences by giving presentations and exhibiting. The library will market itself to the user community as to its available video image file resources.

IV. PILOT PROJECT

In support of the above recommendations, the NOAA Library Pilot Project has been undertaken to demonstrate access to up to 30 digital videos from various NOAA cruises. Access is provided via NOAALINC. This Project depends on cooperation between the NCL and OE staff. Several existing video collections have been considered for the Pilot Project, including the following:

1. Thunder Bay Expedition 2002 cruise collection, which consists of 11 digital videotapes and 7 high-definition digital videotapes, totaling approximately 140 hours.

2. Gulf of Alaska Seamount Exploration cruise collection, which consists of 160 hours of Alvin footage on DV Cam tapes from 13 dives, and 2 one-hour Best-of-Alvin footage on mini DV from each of the two legs of the cruise.

3. Photos and videos taken during the Islands in the Stream 2002 Expedition: Exploring Underwater Oases between July 27 and September 1, 2002, which includes the Johnson-Sea-Link II video footage and still photos from the three missions of the cruise.

The Project objectives are listed below:

1. Provide access to the NOAA OE digital videos online via NOAALINC.

2. Reach and educate the general audience worldwide, including educators and students K-12, academia, scientific communities, etc.

3. Create opportunities for collaboration between OE scientists, IT staff, and NOAA librarians.

For the Project, an existing set of library metadata standards and tools has been used, including the following:

1. MARC 21 metadata standards for input to NOAALINC.

2. Dublin Core metadata standards are also being considered for inclusion on NOAA Library home page or OE portal.

3. OCLC Passport (Windows/client-depending system) and Connexion (Browser-dependent metadata-maker) are used to create metadata records.

Using the above tools and appropriate documentation provided in DV12 (Digital Video 12), DI12 (Digital Image12) forms by the OE PIs, some digital videos and photos—the numbers are growing every day—are available by searching NOAALINC. To quickly retrieve some of the examples, one may put in the search phrase "ocean exploration dv." The metadata records provide information about the whole cruise video collection and provide hot links to the streamed video clips that highlight the entire video footage from the cruise.

V. CONCLUSION

The project staff involved in this project has been very enthusiastic about the opportunity to make the NCL the true archive of NOAA information. This has been a new area for all of us from the Public Services Branch and the Technical Services Branch. One of the main products to come out of this project so far has been some supporting documents created by our metadata librarian; these include a Crosswalk between MARC21/FGDC/Dublin Core Metadata Standards, and Core Elements Required for Metadata Creation for Digital Video Using MARC21, Dublin Core and FGDC Metadata Standards (DV12). These documents are for the cruise PIs to take on the ship to ensure that minimum information and products from the cruise will be received in the NOAA library. Similar documents have also been created for digital images and digital audio, called DI12 and DA12. Twelve core elements were selected to support the creation of metadata records in various metadata standards, including, MARC21, Dublin Core, and FGDC.

While leading the Video/Image/Audio Ocean Exploration Data Management Working Group and being major participants in the development of the NOAA Ocean Exploration Data Management strategy, the NCL team has had the unique opportunity to collaborate with a group of the NOAA scientists, oceanographers, and IT specialists working within the OE Project. Moreover, it has given the group of NOAA librarians an opportunity to receive complete and quality metadata information for input into the NOAALINC catalogs and to provide interested parties access to ocean exploration information.

REFERENCES

[1] *The Nation's Environmental Data: Treasures at Risk: Report to Congress on the Status and Challenges for NOAA's Environmental Data Systems*. Washington, D.C.: U.S. Department of Commerce, National Oceanic and Atmospheric Administration, 2001.

[2] The International Standard, ISO 23950: "Information Retrieval (Z39.50): Application Service Definition and Protocol Specification" and to ANSI/NISO Z39.50.